



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/935,610	08/24/2001	Huiping Li	37112-173148	8980

7590 07/05/2005

Venable
P.O. Box 34385
Washington, DC 20043-9998

EXAMINER

BRIER, JEFFERY A

ART UNIT	PAPER NUMBER
----------	--------------

2672

DATE MAILED: 07/05/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/935,610

Applicant(s)

LI ET AL.

Examiner

Jeffery A. Brier

Art Unit

2672

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 13 April 2005.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-38 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-38 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
 - ☐ Certified copies of the priority documents have been received in Application No. _____.
 - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

Detailed Action

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 4/13/2005 has been entered.

Response to Amendment

2. The amendment filed on 4/13/2005 has been entered. The amendment to paragraph 0034 overcomes the objection to the specification set forth in the previous Final Rejection. The amendment to claims 26-28 do not fully overcome the 35 USC 112 rejection set forth in the previous Final Rejection.

Response to Arguments

3. The arguments filed on 4/13/2005 concerning the Chun article and the Antani article are not persuasive and the amendments made to claims 1 and 35 do not distinguish these claims from the Chun article and the Antani article.

The Chun article focuses on character regions having some fixed colors and sizes, and character regions are densely located in the horizontal direction, as shown in Fig. 2. The colors and shapes are not regular in the background. See section 2. Therefore, Chun recognized the difference between original video with text and text overlayed onto the original video and teaches to one of ordinary skill in the art to discriminate between the original text and the overlayed text.

The Antani article discusses the video having temporal information while the overlayed characters have less temporal information and the overlayed characters are contrasted by a changing background. Abstract and section 4. Text in the original video will more likely have movement from frame to frame. Applicant's arguments made reference to a stop sign example would most likely be part of a moving background while text overlayed onto the video will most likely be stationary. Therefore, Antani recognized the difference between original video with text and text overlayed onto the original video and teaches to one of ordinary skill in the art to discriminate between the original text and the overlayed text.

Claim Rejections - 35 USC § 101

4. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

5. Claims 1-26 and 29-38 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

Art Unit: 2672

Claims 1-25 and 29-38:

These claims do not distinguish from mental steps the operator performed in the prior art systems to determine if text was an overlay onto an original video or a part of the video. In re Prater, 415 F.2d 1393, 162 USPQ 541 (CCPA 1969).

Claim 26:

The “for causing” clause in this claim makes the claim cover both having the computer readable medium outside the computer, it would still be for causing, and inside the computer causing the computer to implement the method. Therefore the scope of this claim covers embodiments where the computer-readable medium containing computer-executable instructions are not a part of the computer, thus, becoming a program per se.

Claim Rejections - 35 USC § 112

6. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

7. Claims 26-28 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The “for causing” clause in these claims make the claims cover both having the computer readable medium outside the computer, it would still be for causing, and inside the computer causing the computer to implement the method. Therefore the scope of this claim covers embodiments where the computer-

readable medium containing computer-executable instructions are not a part of the computer, thus, becoming a program per se. Therefore, these claims are incomplete. The computer of claims 27 and 28 are not concretely connected to the computer readable medium in order to cause the computer to implement the method.

Claim Rejections - 35 USC § 102

8. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

9. Claims 1-3, 22, 23, 26, and 27 are rejected under 35 U.S.C. 102(b) as being anticipated by Byung Tae Chun, Younglae Bae, Tai-Yun Kim, Text Extraction in Videos using Topographical Features of Characters, August 22-25, 1999, IEEE, vol. 2, pages 1126-1130.

This article teaches extracting text from video by two main steps of extracting candidate areas using topographical features and then verifying text is in those areas. Section 3.1 discusses extracting candidate area for text area and section 3.2 discusses verification of candidates of text area. Section 2 discusses character regions having some fixed colors and sizes, and are densely located in the horizontal direction, as shown in Fig. 2. The colors and shapes are not regular in the background. Thus, text in

Art Unit: 2672

the actual video will have movement and will likely have a size different then Chun's algorithm's text size. Therefore, Chun recognized the difference between original video with text and text overlayed onto the original video having original text and teaches to one of ordinary skill in the art to discriminate between the original text and the overlayed text .

A detailed analysis of the claims follows.

Claim 1:

Chun teaches a method of video processing (*See introduction.*) comprising:
extracting a pre-existing overlay present in a video sequence (*See the introduction, paragraph 1 which discusses text appearing in video such as news where is often used to identify people, see figure 2, and to place identifying marks, see the upper left and lower right corners of figure 2.*) said extracting comprising:

detecting at least one potential overlay in the video sequence (*Section 3.1 discusses extracting candidate area for text area.*); and

verifying that each at least one potential overlay is an actual overlay that was previously added to an original video sequence to obtain said video sequence (*Section 3.2 discusses verification of candidates of text area. Section 2 discusses character regions having some fixed colors and sizes, and are densely located in the horizontal direction, as shown in Fig. 2. The colors and shapes are not regular in the background. Thus, text in the actual video will have movement and will likely have a size different then Chun's algorithm's text size. Therefore, Chun recognized the difference between original video with text and text overlayed onto the original video having original text and*

teaches to one of ordinary skill in the art to discriminate between the original text and the overlaid text .).

Claim 2:

Chun teaches the method of claim 1, further comprising the step of post-processing at least one actual overlay to remove extraneous pixels (*Figure 1 shows the post processing step of removing noise. Noise comprises extraneous pixels such as non-character regions inside the character regions, see section 3.3, thus, Chun teaches removing extraneous pixels.*).

Claim 3:

Chun teaches the method of claim 2, wherein said step of post-processing comprises the steps of:

computing a variance for each pixel of the at least one actual overlay (*Section 3.3 discusses removing noise by using Isodata color clustering. The verified actual overlay area is analyzed to determine the color of each pixel to cluster the pixels in the overlay area.*); and

comparing the variance with a threshold to determine whether or not the pixel should be removed as an extraneous pixel (*The size of the color clusters are compared and if they are too small the cluster is removed which removes the pixels forming each cluster.*).

Art Unit: 2672

Claim 22:

Chun teaches the method of Claim 1, wherein said step of detecting comprises the step of:

performing template matching to determine the presence of a potential overlay (*Section 2 and 3.1 discusses using the topological features of characters to determine the presence of a potential overlay. Topological features of characters define a template for each character or groups of characters.*).

Claim 23:

Chun teaches the method of claim 22, wherein said step of detecting further comprises the step of:

determining a template (*The paragraph before section 3 discusses determine n and alpha. The values of n and alpha form a template.*) to be used in said step of performing template matching.

Claim 26:

Chun teaches a computer readable medium containing computer-executable code for causing a computer to implement the method of claim 1 (*Chun discusses using a computer to perform the text extraction in section 4. The discussed Pentium 4 computer using a program written in Visual C++ Ver. 5.0 has the program stored in a computer readable medium, the disk drive and RAM.*).

Claim 27:

Chun teaches a computer system comprising:

a computer (*Chun discusses using a Pentium 4 computer to perform the text extraction in section 4.*); and

a computer readable medium coupled to said computer and containing computer-executable code for causing a computer to implement the method of claim 1 (*Chun discusses using a computer to perform the text extraction in section 4. The discussed Pentium 4 computer using a program written in Visual C++ Ver. 5.0 has the program stored in a computer readable medium, the disk drive and RAM*).

10. Claims 1, 22-27, and 35-38 are rejected under 35 U.S.C. 102(a) as being anticipated by S. Antani, D. Crandall, R. Kasturi, Robust Extraction of Text in Video, Sept 3-7, 2000, IEEE, vol. 1, pages 931-834.

This article teaches detecting static overlays on video by performing a frame to frame comparison of the video. In the section 3, second paragraph at lines 7-11 “artificial caption text” and “scene text occurring naturally in a video frame” is discussed.

The Antani article discusses the video having temporal information while the overlaid characters have less temporal information and the overlaid characters are contrasted by a changing background. Abstract and section 4. Text in the original video will more likely have movement from frame to frame. Applicant's arguments made reference to a stop sign example would most likely be part of a moving background

Art Unit: 2672

while text overlayed onto the video will most likely be stationary. Therefore, Antani recognized the difference between original video with text and text overlayed onto the original video and teaches to one of ordinary skill in the art to discriminate between the original text and the overlayed text.

A detailed analysis of the claims follows.

Claim 1:

Antani teaches a method of video processing (*See introduction.*) comprising:
extracting a pre-existing overlay present in a video sequence (*See the introduction, paragraph 1 second column which discusses text appearing in video.*) said extracting comprising:

detecting at least one potential overlay in the video sequence (*Section 2 discusses three stages, the detection, localization, and segmentation stages. The detection stage detects a potential overlay.*); and

verifying that each at least one potential overlay is an actual overlay that was previously added to an original video sequence to obtain said video sequence (*Section 2 discusses the localization stage which uses many methods to localize the text. Section 2 discusses using many different localization algorithms whose outputs are fused in the spatio-temporal decision fusion module over multiple frames to verify that a potential text is text. Section 2 also discusses using a tracking stage, this would inherently verify the potential text is an actual text. The Abstract and section 4 discusses the video having temporal information while the overlayed characters have less temporal information and the overlayed characters are contrasted by a changing background. Text in the original video will more likely have movement from frame to frame. Applicant's arguments made reference to a stop sign example would most likely*

be part of a moving background while text overlayed onto the video will most likely be stationary. Therefore, Antani recognized the difference between original video with text and text overlayed onto the original video and teaches to one of ordinary skill in the art to discriminate between the original text and the overlayed text.).

Claim 22:

Antani teaches the method of Claim 1, wherein said step of detecting comprises the step of:

performing template matching to determine the presence of a potential overlay (Section 2 discusses the detection of potential overlay in the detection stage which consists of many different localization algorithms whose outputs are fused in the spatio-temporal decision fusion module over multiple frames. In order to determine if text exists then predefined knowledge of the text is compared with the current image to determine if a match exists. Predefined knowledge of the text is a template.).

Claim 23:

Antani teaches the method of claim 22, wherein said step of detecting further comprises the step of:

determining a template to be used in said step of performing template matching (Inherently at some time the templates used by the program were determined.).

Claim 24:

Antani teaches the method of claim 22, wherein said step of verifying comprises the steps of:

performing frame-to-frame correlation of said potential overlay (*Section 2 discusses using many different localization algorithms whose outputs are fused in the the spatio-temporal decision fusion module over multiple frames.*); and

comparing a result of the frame-to-frame correlation with a threshold to determine if the potential overlay is an actual overlay or not (*In order to determine if text exists then predefined knowledge of the text is compared with the current image to determine if a match exists. Predefined knowledge of the text is a template of thresholds.*).

Claim 25:

Antani teaches the method of claim 24, wherein said step of performing frame-to-frame correlation (*See the discussion above for claim 24.*) comprises the steps of:

forming a mean square error over a set of frames from said video sequence, averaged over all of the pixels in said potential overlay (*This claim does not claim a use for the mean square error, thus, a reference that forms a mean square error over a set of frames teaches the claim. This claim does not claim how the mean square error is formed, thus, a reference that inherently forms the error teaches the claim. The specification in paragraph 0039 sets forth a specific formula for calculating the mean square error, however, the claim only broadly claims how the claimed mean square error is calculated. Although the claims are interpreted in light of the specification,*

limitations from the specification are not read into the claims. See In re Van Geuns, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993). The disclosed formula determines the average difference in intensities between a current frame and a subsequent frame. Antani inherently forms the mean square error since Antani in the localization stage fuses over several frames decisions from many localization algorithms which inherently has determined the average difference in intensities between frames in order to determine if text exists.).

Claims 26 and 27:

Inherently the algorithm of Antani is performed by a computer having a computer readable medium containing computer-executable code for causing a computer to implement the claimed steps.

Claim 35:

Antani teaches a method of processing video, comprising:

extracting a pre-existing graphical (*In the section 3 second paragraph lines 7-11*

“artificial caption text” and “scene text occurring naturally in a video frame” is discussed.

Artificial caption text is graphical because graphical includes many objects including text.) overlay present in a video sequence, said extracting comprising:

detecting at least one potential overlay in a the video sequence (*Section 2 discusses three stages, the detection, localization, and segmentation stages. The detection stage detects a potential overlay.*), said detecting comprising the step of:

Art Unit: 2672

performing template matching (*Section 2 discusses the detection of potential overlay in the detection stage which consists of many different localization algorithms whose outputs are fused in the spatio-temporal decision fusion module over multiple frames. In order to determine if text exists then predefined knowledge of the text is compared with the current image to determine if a match exists. Predefined knowledge of the text is a template.*); and

verifying that each at least one potential overlay is an actual overlay that was previously added to an original video sequence to obtain said video sequence (*Section 2 discusses the localization stage which uses many methods to localize the text. Section 2 discusses using many different localization algorithms whose outputs are fused in the the spatio-temporal decision fusion module over multiple frames to verify that a potential text is text. Section 2 also discusses using a tracking stage, this would inherently verify the potential text is an actual text. The Abstract and section 4 discusses the video having temporal information while the overlaid characters have less temporal information and the overlaid characters are contrasted by a changing background. Text in the original video will more likely have movement from frame to frame.*

Applicant's arguments made reference to a stop sign example would most likely be part of a moving background while text overlaid onto the video will most likely be stationary. Therefore, Antani recognized the difference between original video with text and text overlaid onto the original video and teaches to one of ordinary skill in the art to discriminate between the original text and the overlaid text.), said verifying comprising the step of:

performing frame-to-frame correlation of a potential overlay determined by said detecting step (*Section 2 discusses using many different localization algorithms whose outputs are fused in the spatio-temporal decision fusion module over multiple frames.*).

Claim 36:

Antani teaches the method of Claim 35, wherein said step of detecting further comprises the step of:

determining a template to be used in said step of performing template matching (*Inherently at some time the templates used by the program were determined.*).

Claim 37:

Antani teaches the method of Claim 36, wherein said step of determining a template comprises the step of:

performing addition or frame-by-frame subtraction of video frames (*This claim does not define the specifics of the addition of video frames or the frame-by-frame subtraction of video frames. This claim does not state if pixel values are added or frame numbers are added or if as in Antani the results of many frame analyses are fused or added or subtracted.*). This step does not state what function the addition or subtraction performs, thus, the scope of the claim is broad and is met by Antani when a template for detection stage is determined since the claim does not claim when the template is determined and when the addition or subtraction is performed. Therefore in this comprising claim all that is needed is for the reference to teach the claimed steps.

Claim 38:

Antani teaches the method of Claim 36, wherein said step of determining a template comprises the steps of:

segmenting video frames into foreground and background objects (*Text is foreground and video is the background. See the Abstract at the next to last and last sentences.*

Section 1 second paragraph lines 8-9.);

performing correlation tracking to determine if any foreground object remains in the same absolute location in each video frame (*Section 2 discusses using many different localization algorithms whose outputs are fused in the spatio-temporal decision fusion module over multiple frames to verify that a potential text is text. In the last sentence of section 2 the article teaches due to the fact that text lasts over several frames the text may be determined. The Abstract at the last sentence teaches determining if the text is static.*). This step does not state what function the segmenting and correlation tracking performs, thus, the scope of the claim is broad and the claim does not claim when the template is determined and when the segmenting and correlation tracking is performed. Therefore in this comprising claim all that is needed is for the reference to teach the claimed steps.

Prior Art

11. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

The article by Toshio Sato, Takeo Kanade, Ellen K. Hughes, Michael A. Smith titled Video OCR for Digital News Archives (1998) which may be found at <http://citeseer.ist.psu.edu/234543.html> at least in section 2.2.2 teaches determining if the text is an overlay text due to the temporal processing.

The article by Jae-Chang Shim; Dorai, C.; Bolle, R. titled Automatic Text Extraction from Video for Content-Based Annotation and Retrieval, Fourteenth International Conference on Pattern Recognition, 1998. Proceedings. Volume 1, 16-20 Aug. 1998 Page(s): 618 - 620 vol.1 teaches in section 2 extracting superimposed text which is overlaid text.

2. Text Extraction from Video

Text in video appears as either *scene text* or as *superimposed text* [1]. Our system is designed to extract superimposed text and scene text that possesses typical text attributes. We do not assume any prior knowledge about frame resolution, text location, and font styles. Some common characteristics of text are exploited in our algorithm including monochromaticity of individual characters, size restrictions (characters cannot be too small to be read by humans or too big to occupy a large portion of the frame), and horizontal alignment of text.

The article by Ki-Young Jeong; Keechul Jung; Eun Yi Kim; Hang Joon Kim; Neural Network-Based Text Location for News Video Indexing, 1999 International Conference on Image Processing, Volume 3, 24-28 Oct. 1999 Page(s):319 – 323 teaches discriminating between text in the video and overlaid text see figure 7 and associated text. The same article may be found at <http://ailab.kyungpook.ac.kr/vindex/textlocation/textlocation.html> It has additional figures. See figure 8.

Allowable Subject Matter

12. Claims 4-21 would be allowable if rewritten to overcome the rejection(s) under 35 U.S.C. 101, set forth in this Office action and to include all of the limitations of the base claim and any intervening claims. Claims 29-34 would be allowable if rewritten or amended to overcome the rejection(s) under 35 U.S.C. 101, set forth in this Office action. Claim 28 would be allowable if rewritten or amended to overcome the rejection(s) under 35 U.S.C. 101 and 112 second paragraph, set forth in this Office action.

13. The following is a statement of reasons for the indication of allowable subject matter:

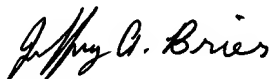
Claims 4-21 and 28-34:

The prior of record fails to teach or suggest detecting the potential overlay by using wavelet decomposition on the video sequence, extracting features based on the

wavelet decomposition, and performing neural network processing on the extracted features.

14. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jeffery A Brier whose telephone number is (571) 272-7656. The examiner can normally be reached on M-F from 7:00 to 3:30. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Michael Razavi, can be reached at (571) 272-7664. The fax phone Number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).


Jeffery A Brier
Primary Examiner
Art Unit 2672